



## SEQUENCE LISTING

<110> Palmer, Michelle A.J.  
Gee, Melissa  
Tillotson, Bonnie  
Chang, Xiao-Jia

<120> Receptor Function Assay for G-Protein Coupled Receptors and Orphan Receptors by Reporter Enzyme Mutant Complementation

<130> 4085-226-27

<140> US 09/654,499  
<141> 2000-09-01

<150> US 60/180,669  
<151> 2000-02-07

<160> 6

<170> FastSEQ for Windows Version 4.0

<210> 1  
<211> 6700  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Nucleotide sequence for pICAST ALC

<221> CDS  
<222> (1457)...(4486)

*E*  
<400> 1  
ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtctca 60  
gggccaagaa cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt 120  
tcctgccccg gctcagggcc aagaacagat ggtccccaga tgcggtccag ccctcagcag 180  
tttctagaga accatcagat gtttccaggg tgccccaagg acctgaaatg accctgtgcc 240  
ttatattgaac taaccaatca gttcgctct cgcttctgtt cgccgcgttc tgctccccga 300  
gctcaataaa agagccccaca acccctcaact cggggcgcga gtcctccgat tgactgagtc 360  
gccccgggtac ccgtgtatcc aataaaacctt ctgcagttg catccgactt gtggtctcgc 420  
tgttccttgg gagggcttcc tctgagtgat tgactaccgg tcaggggggg tctttcattt 480  
gggggcttgt ccgggatcgg gagaccctg cccaggggacc accgaccac caccgggagg 540  
caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgattta 600  
tgccgcctcgc tcggacttag ttagctaact agctctgtat ctggcggacc cgtgggtggaa 660  
ctgacgagtt ctgaacaccc ggccgcaacc ctgggagacg tcccaggggac tttggggggcc 720  
gtttttgtgg cccgacctga ggaagggagt cgatgtggaa tccgaccccg tcaggatatg 780  
tggttcttgtt aggagacgag aacctaataac agtcccgcc tccgtctgaa tttttgtttt 840  
cggtttggaa ccgaagccgc gcgtcttgc tgctgcagca tcgttctgtt ttgtctctgt 900  
ctgactgtgt ttctgttattt gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960  
ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccagtc ggtagatgtc 1020  
aagaagagac gttgggttac cttctgtct gcagaatggc caacctttaa cgtcggatgg 1080  
ccgcgagacg gcaccccttaa ccgagacctc atcacccagg ttaagatcaa ggtctttca 1140  
cctggccccc atggacaccc agaccaggtc ccatacatcg tgacctggga agccttggct 1200  
tttggggccccc ctcctgggtt caagcccttt gtacaccctta agcctccgcc tccctttcct 1260  
ccatccgccc cgtctctccc cttgaaact cctcggtcga ccccgctcg atcctccctt 1320  
tatccagccc tcactccttc tctaggcgcc ggccgctcta gcccattaat acgactca 1380  
atagggcgtatcgaatcagg cttggcgccg ccggatcctt aattaagcgc aattgggagg 1440

tggcggtagc ctcgag atg ggc gtg att acg gat tca ctg gcc gtc gtg gcc 1492  
 Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala  
 1 5 10

cgc acc gat cgc cct tcc caa cag tta cgc agc ctg aat ggc gaa tgg 1540  
 Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp  
 15 20 25

cgc ttt gcc tgg ttt ccg gca cca gaa gcg gtg ccg gaa agc tgg ctg 1588  
 Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu  
 30 35 40

gag tgc gat ctt cct gag gcc gat act gtc gtc ccc tca aac tgg 1636  
 Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp  
 45 50 55 60

cag atg cac ggt tac gat gcg ccc atc tac acc aac gtg acc tat ccc 1684  
 Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro  
 65 70 75

att acg gtc aat ccg ccg ttt gtt ccc acg gag aat ccg acg ggt tgt 1732  
 Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys  
 80 85 90

tac tcg ctc aca ttt aat gtt gat gaa agc tgg cta cag gaa ggc cag 1780  
 Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln  
 95 100 105

acg cga att att ttt gat ggc gtt aac tcg gcg ttt cat ctg tgg tgc 1828  
 Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys  
 110 115 120

aac ggg cgc tgg gtc ggt tac ggc cag gac agt cgt ttg ccg tct gaa 1876  
 Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu  
 125 130 135 140

ttt gac ctg agc gca ttt tta cgc gcc gga gaa aac cgc ctc gcg gtg 1924  
 Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val  
 145 150 155

atg gtg ctg cgc tgg agt gac ggc agt tat ctg gaa gat cag gat atg 1972  
 Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met  
 160 165 170

tgg cgg atg agc ggc att ttc cgt gac gtc tcg ttg ctg cat aaa ccg 2020  
 Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro  
 175 180 185

act aca caa atc agc gat ttc cat gtt gcc act cgc ttt aat gat gat 2068  
 Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp  
 190 195 200

ttc agc cgc gct gta ctg gag gct gaa gtt cag atg tgc ggc gag ttg 2116  
 Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu  
 205 210 215 220

---

cgt gac tac cta cgg gta aca gtt tct tta tgg cag ggt gaa acg cag 2164  
 Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln  
 225 230 235

gtc gcc agc ggc acc gcg cct ttc ggc ggt gaa att atc gat gag cgt	2212
Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg	
240 245 250	
ggt ggt tat gcc gat cgc gtc aca cta cgt ctg aac gtc gaa aac ccg	2260
Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro	
255 260 265	
aaa ctg tgg agc gcc gaa atc ccg aat ctc tat cgt gcg gtg gtt gaa	2308
Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu	
270 275 280	
ctg cac acc gcc gac ggc acg ctg att gaa gca gaa gcc tgc gat gtc	2356
Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val	
285 290 295 300	
ggt ttc cgc gag gtg cggtt gaa aat ggt ctg ctg ctg ctg aac ggc	2404
Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu Leu Leu Asn Gly	
305 310 315	
aag ccg ttg ctg att cga ggc gtt aac cgt cac gag cat cat cct ctg	2452
Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu His His Pro Leu	
320 325 330	
cat ggt cag gtc atg gat gag cag acg atg gtg cag gat atc ctg ctg	2500
His Gly Gln Val Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu	
335 340 345	
atg aag cag aac aac ttt aac gcc gtg cgc tgt tcg cat tat ccg aac	2548
Met Lys Gln Asn Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn	
350 355 360	
cat ccg ctg tgg tac acg ctg tgc gac cgc tac ggc ctg tat gtg gtg	2596
His Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val	
365 370 375 380	
gat gaa gcc aat att gaa acc cac ggc atg gtg cca atg aat cgt ctg	2644
Asp Glu Ala Asn Ile Glu Thr His Gly Met Val Pro Met Asn Arg Leu	
385 390 395	
acc gat gat ccg cgc tgg cta ccg gcg atg agc gaa cgc gta acg cga	2692
Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg	
400 405 410	
atg gtg cag cgc gat cgt aat cac ccg agt gtg atc atc tgg tcg ctg	2740
Met Val Gln Arg Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu	
415 420 425	
ggg aat gaa tca ggc cac ggc gct aat cac gac ggc ctg tat cgc tgg	2788
Gly Asn Glu Ser Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp	
430 435 440	
atc aaa tct gtc gat cct tcc cgc ccg gtg cag tat gaa ggc ggc gga	2836
Ile Lys Ser Val Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Gly	
445 450 455 460	
gcc gac acc acg gcc acc gat att att tgc ccg atg tac gcg cgc gtg	2884
Ala Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val	
465 470 475	

Page 3

gat gaa gac cag ccc ttc ccg gct gtg ccg aaa tgg tcc atc aaa aaa 2932  
 Asp Glu Asp Gln Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys  
 480 485 490

tgg ctt tcg cta cct gga gag acg cgc ccg ctg atc ctt tgc gaa tac 2980  
 Trp Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr  
 495 500 505

gcc cac gcg atg ggt aac agt ctt ggc ggt ttc gct aaa tac tgg cag 3028  
 Ala His Ala Met Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln  
 510 515 520

gcg ttt cgt cag tat ccc cgt tta cag ggc ggc ttc gtc tgg gac tgg 3076  
 Ala Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp  
 525 530 535 540

gtg gat cag tcg ctg att aaa tat gat gaa aac ggc aac ccg tgg tcg 3124  
 Val Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser  
 545 550 555

gct tac ggc ggt gat ttt ggc gat acg ccg aac gat cgc cag ttc tgt 3172  
 Ala Tyr Gly Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys  
 560 565 570

atg aac ggt ctg gtc ttt gcc gac cgc acg ccg cat cca gcg ctg acg 3220  
 Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr  
 575 580 585

gaa gca aaa cac cag cag ttt ttc cag ttc cgt tta tcc ggg caa 3268  
 Glu Ala Lys His Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln  
 590 595 600

acc atc gaa gtg acc agc gaa tac ctg ttc cgt cat agc gat aac gag 3316  
 Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu  
 605 610 615 620

ctc ctg cac tgg atg gtg gcg ctg gat ggt aag ccg ctg gca agc ggt 3364  
 Leu Leu His Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly  
 625 630 635

gaa gtg cct ctg gat gtc gct cca caa ggt aaa cag ttg att gaa ctg 3412  
 Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu  
 640 645 650

cct gaa cta ccg cag ccg gag agc gcc ggg caa ctc tgg ctc aca gta 3460  
 Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val  
 655 660 665

cgc gta gtg caa ccg aac gcg acc gca tgg tca gaa gcc ggg cac atc 3508  
 Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile  
 670 675 680

agc gcc tgg cag cag tgg cgt ctg gcg gaa aac ctc agt gtg acg ctc 3556  
 Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu  
 685 690 695 700

ccc gcc gcg tcc cac gcc atc ccg cat ctg acc acc agc gaa atg gat 3604  
 Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp  
 705 710 715

C  
 1  
 CNT

ttt tgc atc gag ctg ggt aat aag cgt tgg caa ttt aac cgc cag tca 3652  
 Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser  
 720 725 730  
  
 ggc ttt ctt tca cag atg tgg att ggc gat aaa aaa caa ctg ctg acg 3700  
 Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr  
 735 740 745  
  
 ccg ctg cgc gat cag ttc acc cgt gca ccg ctg gat aac gac att ggc 3748  
 Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly  
 750 755 760  
  
 gta agt gaa gcg acc cgc att gac cct aac gcc tgg gtc gaa cgc tgg 3796  
 Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp  
 765 770 775 780  
  
 aag gcg ggc cat tac cag gcc gaa gca gcg ttg ttg cag tgc acg 3844  
 Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr  
 785 790 795  
  
 gca gat aca ctt gct gat gcg gtg ctg att acg acc gct cac gcg tgg 3892  
 Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp  
 800 805 810  
  
 cag cat cag ggg aaa acc tta ttt atc agc cgg aaa acc tac cgg att 3940  
 Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile  
 815 820 825  
  
 gat ggt agt ggt caa atg gcg att acc gtt gat gtt gaa gtg gcg agc 3988  
 Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser  
 830 835 840  
  
 gat aca ccg cat ccg gcg cgg att ggc ctg aac tgc cag ctg gcg cag 4036  
 Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln  
 845 850 855 860  
  
 gta gca gag cgg gta aac tgg ctc gga tta ggg ccg caa gaa aac tat 4084  
 Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr  
 865 870 875  
  
 ccc gac cgc ctt act gcc gcc tgt ttt gac cgc tgg gat ctg cca ttg 4132  
 Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu  
 880 885 890  
  
 tca gac atg tat acc ccg tac gtc ttc ccg agc gaa aac ggt ctg cgc 4180  
 Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg  
 895 900 905  
  
 tgc ggg acg cgc gaa ttg aat tat ggc cca cac cag tgg cgc ggc gac 4228  
 Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp  
 910 915 920  
  
 ttc cag ttc aac atc agc cgc tac agt caa cag caa ctg atg gaa acc 4276  
 Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Leu Met Glu Thr  
 925 930 935 940  
  
 agc cat cgc cat ctg ctg cac gcg gaa gca ggc aca tgg ctg aat atc 4324  
 Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile  
 945 950 955

gac ggt ttc cat atg ggg att ggt ggc gac gac tcc tgg agc ccg tca Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser 960 965 970	4372
gta tcg gcg gaa ttc cag ctg agc gcc ggt cgc tac cat tac cag ttg Val Ser Ala Glu Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu 975 980 985	4420
gtc tgg tgt caa aaa aga tct gac tat aaa gat gag gac ctc gac cat Val Trp Cys Gln Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His 990 995 1000	4468
cat cat cat cac cggttaataatagg tagataagtg actgattaga His His His His Arg 1005 1010	4516
tgcatggatc cctcgaccaa ttccggttat ttccacat attggcgctc ttggcaatg tgaggggccg gaaacctggc cctgtttct tgacgagcat tcctagggggt cttccctc tcgccaagg aatgcaggat ctgttgaatg tcgtgaagga agcagttcct ctggaaagct cttgaagaca aacaacgtct gtacgaccc ttgcaggca gcggaaacccc ccacctggcg acagggtgcct ctgcggccaa aagccacgtg tataagatac acctgcaaaag gcggcacaac cccagtgcct cttgtgagt tggatagtt tggaaagagt caaatggctc tcctcaagcg tattcaacaa ggggctgaag gatgcccaga aggtacccca ttgtatggg tctgatctgg ggcctcggtg cacatgctt acatgtgtt agtcgaggt aaaaaacgtc taggcffff gaaccacggg gacgtggtt tccttgaaa aacacgtga taataccatg attgaacaag atggatgcg cgcagggttc cggccgcgtt gggtggagag gctattcgcc tatgactgg cacaacagac aatcggtgc tctgatgcc cctgtttccg gctgtttccg gctgtcagcg cggttcttt tgcataagacc gacctgtccg gtgcctgtaa tgaactgcag gacgagggcag cgccgtatc gtggctggcc acgacggcg ttccttgcc agctgtgctc gacgttgtca ctgaaggccgg aaggactgg ctgttattgg tgcgttgc tggcggaggat ctcctgtcat ctcaccttgc tcctgcccgg aaagtatcca cgttgc tgcataatggc cggctgcata cgcttgc tgcataatggc ccattgc accaaggcga acatgcgc tggcggaggat gtactcgat ggaaggccgg tttgtgc tgcataatggc cggctgcata tcgcgcgc cgaactgttc gcccggctca tgcgttgc tggcggaggat tcgtgaccca tggcgtatgcc tgcgttgc tggcggaggat gattcatcg tgcgttgc tggcggaggat tgcataatggc cggctgcata cccggtat tgcataatggc tggcggaggat tgcataatggc cggctgcata gtatcgccgc tcccgatttc cagcgcatcg tgcgttgc tggcggaggat gagcgggact ctgggggttc catcgataaa tgcataatggc cggctgcata ggggggaaatg aaagacccca cctgttaggtt tgcataatggc cggctgcata caaggcatgg aaaaatatacat aactgagaat tgcataatggc cggctgcata tggaaacactg gatataatggc caaacaggat tgcataatggc cggctgcata aggcggcaga acagatggaa cagctgaata tgcataatggc cggctgcata ggggggaaatg aaaaatatacat aactgagaat tgcataatggc cggctgcata ttcctgcccc ggctcgaggc caagaacaga tgcataatggc cggctgcata gtttcttagag aaccatcaga tgcgttgc tggcggaggat cttatttgc tgcataatggc tggcggaggat tgcataatggc cggctgcata cttatttgc tgcataatggc tggcggaggat tgcataatggc cggctgcata agctcaataa aagagcccac aaccctcact tgcataatggc cggctgcata cgcccggtt cccgtgtatc caataaaaccc tgcataatggc cggctgcata ctgttgc tgcataatggc tggcggaggat tgcataatggc cggctgcata catgcagcat gatcaaaaat tgcataatggc tggcggaggat tgcataatggc atagttgc tgcataatggc tggcggaggat tgcataatggc cggctgcata tccgcttgc tgcataatggc tggcggaggat tgcataatggc cggctgcata gctcactcaa aggccgtat acgg tgcataatggc cggctgcata	4576 4636 4696 4756 4816 4876 4936 4996 5056 5116 5176 5236 5296 5356 5416 5476 5536 5596 5656 5716 5776 5836 5896 5956 6016 6076 6136 6196 6256 6316 6376 6436 6496 6556 6616 6676 6700

<210> 2  
<211> 1010  
<212> PRT  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Amino acid sequence for pICAST ALC

&lt;400&gt; 2

Met Gly Val Ile Thr Asp Ser Leu Ala Val Val Ala Arg Thr Asp Arg  
 1 5 10 15  
 Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp  
 20 25 30  
 Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu Glu Cys Asp Leu  
 35 40 45  
 Pro Glu Ala Asp Thr Val Val Pro Ser Asn Trp Gln Met His Gly  
 50 55 60  
 Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro Ile Thr Val Asn  
 65 70 75 80  
 Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys Tyr Ser Leu Thr  
 85 90 95  
 Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln Thr Arg Ile Ile  
 100 105 110  
 Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys Asn Gly Arg Trp  
 115 120 125  
 Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu Phe Asp Leu Ser  
 130 135 140  
 Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val Met Val Leu Arg  
 145 150 155 160  
 Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met Trp Arg Met Ser  
 165 170 175  
 Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro Thr Thr Gln Ile  
 180 185 190  
 Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp Phe Ser Arg Ala  
 195 200 205  
 Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu Arg Asp Tyr Leu  
 210 215 220  
 Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln Val Ala Ser Gly  
 225 230 235 240  
 Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg Gly Gly Tyr Ala  
 245 250 255  
 Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro Lys Leu Trp Ser  
 260 265 270  
 Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu Leu His Thr Ala  
 275 280 285  
 Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val Gly Phe Arg Glu  
 290 295 300  
 Val Arg Ile Glu Asn Gly Leu Leu Leu Asn Gly Lys Pro Leu Leu  
 305 310 315 320  
 Ile Arg Gly Val Asn Arg His Glu His His Pro Leu His Gly Gln Val  
 325 330 335  
 Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu Met Lys Gln Asn  
 340 345 350  
 Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn His Pro Leu Trp  
 355 360 365  
 Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val Asp Glu Ala Asn  
 370 375 380  
 Ile Glu Thr His Gly Met Val Pro Met Asn Arg Leu Thr Asp Asp Pro  
 385 390 395 400  
 Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg Met Val Gln Arg  
 405 410 415  
 Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu Gly Asn Glu Ser  
 420 425 430  
 Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp Ile Lys Ser Val  
 435 440 445  
 Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Ala Asp Thr Thr  
 450 455 460

 G1  
 G2  
 Cont

Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val Asp Glu Asp Gln  
 465 470 475 480  
 Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys Trp Leu Ser Leu  
 485 490 495  
 Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr Ala His Ala Met  
 500 505 510  
 Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln Ala Phe Arg Gln  
 515 520 525  
 Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp Val Asp Gln Ser  
 530 535 540  
 Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly  
 545 550 555 560  
 Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys Met Asn Gly Leu  
 565 570 575  
 Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr Glu Ala Lys His  
 580 585 590  
 Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln Thr Ile Glu Val  
 595 600 605  
 Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu Leu Leu His Trp  
 610 615 620  
 Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly Glu Val Pro Leu  
 625 630 635 640  
 Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu Pro Glu Leu Pro  
 645 650 655  
 Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val Arg Val Val Gln  
 660 665 670  
 Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile Ser Ala Trp Gln  
 675 680 685  
 Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu Pro Ala Ala Ser  
 690 695 700  
 His Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp Phe Cys Ile Glu  
 705 710 715 720  
 Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser Gly Phe Leu Ser  
 725 730 735  
 Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr Pro Leu Arg Asp  
 740 745 750  
 Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly Val Ser Glu Ala  
 755 760 765  
 Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp Lys Ala Ala Gly  
 770 775 780  
 His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu  
 785 790 795 800  
 Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp Gln His Gln Gly  
 805 810 815  
 Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly  
 820 825 830  
 Gln Met Ala Ile Thr Val Asp Val Glu Val Ala Ser Asp Thr Pro His  
 835 840 845  
 Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln Val Ala Glu Arg  
 850 855 860  
 Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr Pro Asp Arg Leu  
 865 870 875 880  
 Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu Ser Asp Met Tyr  
 885 890 895  
 Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg Cys Gly Thr Arg  
 900 905 910  
 Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly Asp Phe Gln Phe Asn  
 915 920 925  
 Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met Glu Thr Ser His Arg His  
 930 935 940  
 Leu Leu His Ala Glu Glu Gly Thr Trp Leu Asn Ile Asp Gly Phe His  
 945 950 955 960

Met Gly Ile Gly Gly Asp Asp Ser Trp Ser Pro Ser Val Ser Ala Glu  
 965 970 975  
 Phe Gln Leu Ser Ala Gly Arg Tyr His Tyr Gln Leu Val Trp Cys Gln  
 980 985 990  
 Lys Arg Ser Asp Tyr Lys Asp Glu Asp Leu Asp His His His His His  
 995 1000 1005  
 His Arg  
 1010

<210> 3  
 <211> 8518  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Nucleotide sequence for pICAST ALN

<400> 3  
 ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca 60  
 gggccaagaa cagatggAAC agctgaatat gggccaaaca ggatatctgt gttaaggca 120  
 tcctgccccg gctcaggGCC aagaacagat ggtccccaga tgcggtccag ccctcagcag 180  
 tttctagaga accatcagat gtttccaggg tgcccaagg acctgaaatg accctgtgcc 240  
 ttatTTgaac taaccaatca gttcgcttct cgcttctgtt cgcgcgttc tgctccccga 300  
 gctcaataaa agagccccaca acccctcact cggggcgcCA gtcctccgat tgactgagtc 360  
 gcccgggtac ccgtgtatcc aataaaccct cttgcagttg catccgactt gtggtctcgc 420  
 tgttccttgg gagggtctcc tctgagtgat tgactacccg tcagggggg tcttcattt 480  
 gggggctcgt ccgggatcgg gagacccctg cccaggAAC accgacccac caccggggagg 540  
 caagctggcc agcaacttat ctgtgtctgt ccgattgtct agtgtctatg actgatttt 600  
 tgcgcctgCG tcggtaactg ttagctaact agctctgtat ctggcggacc cgtggtgaa 660  
 ctgacgaggat ctgaacaccc ggccgaacc ctgggagacg tcccaggac tttggggggc 720  
 gtttttggg cccgacctga ggaaggaggat cgatgtggaa tccgaccccg tcaggatatg 780  
 tggttctgtt aggagacgag aacctaaaac agttccgcC tccgtctgaa tttttgttt 840  
 cggtttggaa ccgaagccgc gctgttgc tgcgtcagca tcgttctgtg ttgtctctgt 900  
 ctgactgtgt ttctgtatTT gtctgaaaat tagggccaga ctgttaccac tcccttaagt 960  
 ttgaccttag gtaactggaa agatgtcgag cggctcgctc acaaccatgc gtagatgtc 1020  
 aagaagagac gttgggttac ctctgtctc gcaaaatggc caacctttaa cgtcgatgg 1080  
 cccgcagacg gcacctttaa ccgagacctc atcacccagg ttaagatcaa gtcctttca 1140  
 cctggccccc atggacaccc agaccaggTC ccctacatcg tgacctggg agccttggt 1200  
 tttgacccccc ctccctgggt caagccctt gtacacccta agcctccgc tcctcttct 1260  
 ccatccgccc cgtctctccc cttgaacct cctcggtcga ccccgctcg atcctccctt 1320  
 tatccagccc tcactccttc tctaggcgcc ggcgcctca gcccattaaat acgactca 1380  
 atagggcgtat tcgaacacca tgcaccatca tcatcatcac gtcgactata aagatgagga 1440  
 cctcgagatg ggcgtgatta cggattcact ggccgtcgT gcccgcaccg atcgcccttc 1500  
 ccaacagttt cgcagcctga atggcgaatg gctgttgc tgggttccgg caccagaagc 1560  
 ggtgcggaa agctggctgg agtgcgtatct tcctgaggcc gatactgtcg tcgtccccc 1620  
 aaactggcag atgcacggTT acgatgcGCC catctacacc aacgtgaccc atcccattac 1680  
 ggtcaatccg ccgttggTC ccacggagaa tccgacgggt tgtaactcgc tcacattaa 1740  
 tggtgatgaa agctggctac aggaaggcCA gacgcgaatt atttttgtatg gcgttaactc 1800  
 ggcgtttcat ctgtggtgca acgggcgtg ggtcggttac gcccaggaca gtcgtttgcc 1860  
 gtctgaattt gacctgagcg cattttacat cgcggagaa aaccgcctcg cggtgatgt 1920  
 gctgcgtgg agtgcacggca gttatctgg agatcaggat atgtggcggA tgagcggcat 1980  
 tttccgtgac gtctcgTTgc tgcataaaacc gactacacaa atcagcgatTT tcctatgttgc 2040  
 cactcgctt aatgtatgatt tcagccgcC tgcgtggag gctgaagtTC agatgtgcgg 2100  
 cgagttcggt gactacctac gggtaacagt ttctttatgg cagggtaaaa cgcaggTCgc 2160  
 cagccggacc cgcgcTTTcg cgggtgaat tatcgatgag cgtgggtggT atgcgcgtc 2220  
 cgtcacacta cgtctgaacg tcgaaaaccc gaaactgtgg agcgcggaaa tcccgaaatc 2280  
 ctatcgTgcG gtgggtgaac tgcacaccgc cgcacggcAc ctgattgaag cagaaggctg 2340  
 cgatgtcggt ttccgcgagg tgcggatgA aatggTctg ctgctgtca acggcaaggc 2400  
 gttgctgatt cgaggcgtta accgtcacga gcatcatcct ctgcgtggc aggtcatgga 2460  
 tgagcagacg atggTgcagg atatcctgt gatgaaggcag aacaacttta acggcgtgcg 2520  
 ctgttcgcattatccgaacc atccgctgtg gtacacgcgt acggcctgtA 2580

6  
 Cont

tggtggat gaagccaata ttgaaaccca cggcatggtg ccaatgaatc gtctgaccga 2640  
tgcgtccgc tgcttacccgg cgatgagcga acgcgtaaacg cgaatggtgc agcgcgatcg 2700  
taatcccccg agtgtatca tctggtcgct ggggaatgaa tcaggccacg ggcctaatac 2760  
cgacgcgctg tatcgctggc tcaaatctgt cgatccttc cgccccgggtc agtatgaagg 2820  
cgccggagcc gacaccacgg ccaccgatata tatttgcctc atgtacgcgc ggcgtggatga 2880  
agaccagccc ttccggctg tgccgaaatg gtccatcaaa aatggctt cgctacctgg 2940  
agagacgcgc cogctgatcc tttgcaata cgcccacgcg atggtaaca gtcttggcgg 3000  
tttcgctaaa tactggcagg cgtttcgtca gtatccccgt ttacaggcgc gcttcgtctg 3060  
ggactgggtg gatcagtgc tgattaaata tgatgaaaac ggcacccgt ggtcgctta 3120  
cgccgggtat ttggcgata cgccgaaacga tcgccagttc tggatgaaacg gtctggctt 3180  
tgccgaccgc aecgcgcac tgcgtcgtac ggaagcaaaa caccagcagc agttttcca 3240  
gttccgttta tccggcaaa ccatcgaagt gaccagcga tacctgttcc gtcatacgca 3300  
taacgagctc ctgcactggc tggggcgct ggtggtaag ccgctggcaa gcggtgaagt 3360  
gcctctggat gtcgctccac aaggtaaaca gttgattgaa ctgcctgaac taccgcagcc 3420  
ggagagcgcc gggcaactct ggctcacagt acgcgtatgt caaccgaacg cgaccgcatg 3480  
gtcagaagcc gggcacatca ggccttggca gcaagtggcgt ctggcgaaa acctcaagtgt 3540  
gacgctcccc gccgcgtccc acgcacatccc gcatctgacc accagcgaaa tggatttttg 3600  
catcgagctg gtaataaagc gttgcaatt taaccgcag tcaggcttcc ttacacagat 3660  
gtggattggc gataaaaaaac aactgctgac gccgctgcgc gatcagtta cccgtgcacc 3720  
gctggataac gacattggc taagtgaagc gaaaggcatt gaccctaaacg cctgggtcga 3780  
acgctgaaag gccggggggc attaccaggc cgaagcagcg ttgttgcgt gcacggcaga 3840  
tacactgtct gatcggtgc tgattacgac cgctcaacgc tggcagcatc agggggaaaac 3900  
cttattttatc agccggaaaaa cctaccggat tgatggtagt ggtcaaatgg cgattaccgt 3960  
tgatgtgaa gtggcgagcg atacaccgc tccggcgcgg attggctgaa actgcacgt 4020  
ggcgcgagta gagaaggcgaaa taaactgggt cggattaggc cgcgaagaaa actatccga 4080  
ccgcctact gccgcctgtt ttgaccgcgt gcatctggca ttgtcagaca tgcataccccc 4140  
gtacgtcttc ccgagcgaaa acggctgcgc acaccagtgg cgccggcact tccagttaa 4200  
acaccagtgg cgccggcact ttccatcatc tgctgcacgc ggaagaaggc acatggctga atatcgacgg 4260  
ggaaaccaggc catcgccatc ttccatcatg gggattgggt gcgacgactc ctggagcccg tcagttatcg 4320  
gctgagcgcc ggtcgctacc attaccagtt ggtctgggt caaaaaagat ctggaggtgg 4380  
tggcagcagg cctggcgcgc ccggatcctt ataagtgact gattagatgc attgatccct gaccaattc cggttatttt ccaccatatt 4440  
ataaggact gatggatgc ggtggcttt ggcaatgtga gggccggaa acctggccct gtcttcttga cgagcattcc 4500  
taggggtctt tcccctctcg ccaaaggaaat gcaaggctcg ttgaatgtcg tgaaggaagc 4560  
agttccctcg gaagcttctt gaagacaac aacgtctgtt ggcaccctt gcagggcagcg 4620  
gaacccccc cctggcgaca ggtgcctctg tgcaaaaggcg gaccaatcc acctggccct gtcttcttga 4680  
tgcaaaaggcg gcacaacccc agtggcacgt gcaaggctcg ttgaatgtcg tgaaggaagc 4740  
atggctctcc tcaagctat tcaacaaggg aacgtctgtt ggcaccctt gcagggcagcg 4800  
tatggatct gatctggggc ctcgtgcac aaacgtcttag gccccccgaa ccacggggac tgtaggttgg atagttgtgg 4860  
aaacgtcttag gccccccgaa ccacggggac taccatgatt gaacaagat gattgcacgc gctgaaggat gcccagaagg 4920  
attcggttat gactgggcac aacagacaat gtcagcgcag gggcccccgg ttcttttgc atgcttaca tggatgtttagt 4980  
gtcagcgcag gggcccccgg ttcttttgc actgcaggac gaggcagcgc ggctatcgtg tggttttcc tttgaaaaac 5040  
tacatggatctc ctgtcatctc accttgcgtcc tgccgatctc ggtggccacg acggggcgttc aggttctccg gccgcttggg 5100  
aatcgccgg ctgcatacgc ttgatccggc caagaccgac ctgtccgggtg gctggccacg acggggcgttc tggagaggct 5160  
tcgcatacgag cgagcacgtt ctcgtatggc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5220  
cgaaagacgt caggggctcg cgccacgcga tgccgatctc ggtggccacg acggggcgttc tggatgtttagt 5280  
cgacggcgag gatctcgatc tgacccatgg aatcgccgg ctgtccgggtg gctggccacg acggggcgttc tgccgatctc 5340  
aaatggccgc ttttctggat tcatcgactg ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5400  
ggacatagcg ttggctacc tcatcgactg ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5460  
cttcctcgat ctttacggta tgcggctcc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5520  
tcttgcggat ttcttcgtat cgggactctg ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5580  
tttagtctcc agaaaaagggg gggaaatgaa gacccacatc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5640  
taagtaacgc cattttgcata ggcatacgaa gatcgatctc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5700  
tcaaggatcg gacacatcgaa aacagactcg gatcgatctc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5760  
agttccctgat cccgctcagg gccaagaaca tcatcgactg ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5820  
atatctgtgg taagcagttc ctgcggccgc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5880  
cggtccagcc ctcaacatcg tctagagaac gatcgatctc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 5940  
ctqaaatqac cctgtqcctt atttqaacta gacccacatc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 6000  
aatacatatac tggaaataga gaaatcgatc tggatgtttagt tgccgatctc 6060  
tatggggccaa acaggatatac gatcgatctc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 6120  
atatctgtgg taagcagttc ctgcggccgc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 6180  
cggtccagcc ctcaacatcg tctagagaac gatcgatctc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 6240  
ctqaaatqac cctgtqcctt atttqaacta gacccacatc ggtggccacg acggggcgttc tggatgtttagt tgccgatctc 6300

cgcgcttctg ctccccgagc tcaataaaag agcccacaac ccctcactcg gggcgccagt 6360  
 cctccgattg actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tcgagttgca 6420  
 tccgacttgt ggtctcgctg ttccctggga gggctcctc tgagtgattg actaccgc 6480  
 agcgggggtc tttcattcat gcagcatgtt tcaaaattaa tttgggtttt tttcttaagt 6540  
 atttacatta aatgccata gttgcattaa tgaatcgcc aacgcgcggg gagaggc 6600  
 ttgcgtattg ggccttcc tcactgactc gctgcgtcg gtcgttcggc 6660  
 tgccgcgagc ggtatcagct cactcaaagg cgtaataacg gttatccaca gaatcagggg 6720  
 ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggcaggaaac ctgtaaaaagg 6780  
 cccgcttgc ggcgttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 6840  
 gctcaagtca gaggtggcga aacccgacag gactataaaag ataccaggcg tttcccccgt 6900  
 gaagctccct cgtgcgtct cctgtccga ccctgcccgt taccggatac ctgtccgc 6960  
 ttctccctc gggaaagctg ggcgttttc atagctcagct ctgttaggtat ctcagttcgg 7020  
 ttaggtctg tcgctccaag ctgggctgtg tgcaagcaacc ccccggtcag cccgaccgct 7080  
 ggccttatac cggttaactat cgctttagt ccaacccgtt aagacacgac ttatcgccac 7140  
 tggcagcagc cactggtaac aggattagca gagcggagta ttagggcggt gctacagagt 7200  
 tcttgaagtg gtggcttaac tacggctaca ctagaagaac agtatttggt atctgcgc 7260  
 tgctgaagcc agttacctc ggaaaaaagag ttgttagctc ttgatccggc aaacaaacca 7320  
 ccgcttggtag cggtggttt ttgttttgc aagcagcagat tacgcgcaga aaaaaaggat 7380  
 ctcagaaga tccttgc ttttctacgg ggctgacgc tcagtgaaac gaaaactcac 7440  
 gttaaggat ttgttgc ttttgc ttttgc ttttgc ttttgc 7500  
 cgcaaatcaa tctaaagtat atatgat ttttgc ttttgc ttttgc 7560  
 agtggggcac ctatctcgc gatctgtcta ttctgttcat ccatagttgc ctgactcccc 7620  
 gtcgtgtaga taactacgt acgggaggc ttaccatctg gccccagtgc tgcaatgata 7680  
 cccgcgagacc cacgttcacc ggctccagat ttatcagcaa taaaccagcc agccggaaagg 7740  
 gcccggcga gaagtggtcc tgcaacttta tccgcctcca tccagtttat taattttgc 7800  
 cgggaagctt gagaatgtt ttcggcaattt aatagttgc gcaacgttgc tgccattgt 7860  
 acaggcatcg tgggtgc acgggaggc ttaccatctg gccccagtgc tgcaatgata 7920  
 cgatcaaggc gagttacatg atccccatg ttgtcaaaa aagcggttag ctccctcgt 7980  
 cctccgatcg ttgtcagaag taagttggcc gcaatgttcat ttttgc ttttgc 8040  
 ctgcataatt ctcttactgt catgcattcc gtaagatgtt ttttgc ttttgc 8100  
 tcaaccaagt catttcgaga atagttgtatc cggcgaccga gttgtcttgc cccggcg 8160  
 atacgggata ataccgcgc acatagcaga actttaaaag tgctcatcat tggaaaacgt 8220  
 tcttcggggc gaaaacttc aaggatctt ccgtgttgc gatccagttc gatgtaaacc 8280  
 actcgtgcac ccaactgtatc ttctgttca ccagcgttgc tgggtgagca 8340  
 aaaacaggaa ggcaaaatgc cgcaaaaaaag ggaataaggcgacacggaa atgttgaata 8400  
 ctctatacttc tccttttca atattatgtt agcatttac aggttatttgc tctcatgagc 8460  
 ggatacatat ttgtatgtt ttagaaaaat aaacaaatag gggttccgcg cacatttc 8518

*9/*  
*6*  
*Cont*  
 <210> 4  
 <211> 8175  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Nucleotide sequence for pICAST OMC

<400> 4  
 ctgcagcctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca 60  
 gggccaagaa cagatggaaac agctgaatat gggccaaaca ggatatctgt ggtaaggc 120  
 tcctgccccg gctcaggccc aagaacagat ggtccccaga tgccgtccag ccctcagc 180  
 ttcttagaga accatcagat gttccaggg tgccccaaagg acctgaaatg accctgtcc 240  
 ttatgttgc acatgttgc gttcgttctt cgcttctgtt cgccgcgttc tgctccccg 300  
 gctcaataaa agagcccaca acccctcaact cggggccca gtcctccgt tgactgagtc 360  
 gcccgggtac ccgtgtatcc aataaaaccct ttcgtcgttgc catccgactt gtgggtctc 420  
 ttttccttgg gaggtctcc ttcgtacttgc tgacttccgc tcaagggggg ttttgcattt 480  
 gggggctcg tccggatcg gagaatccctg cccaggacc accgaccac caccgggg 540  
 caagctggcc agcaattat ctgtgtctgt ccgttgc ttttgc ttttgc ttttgc 600  
 tgccctcg tccgtacttgc ttagtcaact agctctgtat ctggcgacc cgtgggtggaa 660  
 ctgacgatgt ctgaacaccc ggccgcacc ctgggagacg tcccgaggac tttggggggc 720  
 gttttgtgg cccgacacttca ggaaggaggat cgatgtggaa tccgaccctg tcaggatatg 780  
 tgggtctgtt aggagacgag aacctaaaac agtcccgcc tccgtctgaa ttttgc 840  
 cgggttggaa ccgaagccgc gctgttgc tgctgcagca tcgttctgtt tggtctctgt 900

ctgactgtgt	ttctgttatt	gtctgaaaat	tagggccaga	ctgttaccac	ccctctaagt	960
ttgacccctag	gtaactggaa	agatgtcgag	cggctcgctc	acaaccagtc	ggtagatgtc	1020
aagaagagac	gttgggttac	cttctgtct	gcagaatggc	caacctttaa	cgtcgatgg	1080
cccgagacg	gcacctttaa	ccgagacctc	atcaccagg	ttaagatcaa	ggtctttca	1140
cctggccgc	atggacaccc	agaccaggc	ccctacatcg	tgacctggga	agcctggct	1200
tttgacccccc	ctccctgggt	caagccctt	gtacacccta	agctccgccc	tcctcttcct	1260
ccatccgccc	cgtctctccc	ccttgaacct	cctcggtcga	ccccgcctcg	atcctccctt	1320
tatccagccc	tcactccttc	tctaggcgcc	ggccgctcta	gccattaaat	acgactcaact	1380
ataggcgat	tgcgaatcagg	ccttggcgcg	ccggatcctt	aattaagcgc	aattgggagg	1440
tggcggttagc	ctcgagatgg	gcgtgattac	ggattcaactg	gccgtcggtt	tacaacgtcg	1500
tgactggaa	aaccctggcg	ttacccaact	taatcgccctt	gcagcacatc	ccccttcgc	1560
cagctggcgt	aatagcgaag	aggcccgcac	cgatcgccct	tcccaacagt	tacgcagcct	1620
gaatggcgaa	tggcgctttg	cctggttcc	ggcaccagaa	gcggtgccgg	aaagctggct	1680
ggagtgcgtat	cttcctgagg	ccgatactgt	cgtcgtcccc	tcaaactggc	agatgcacgg	1740
ttacgatgctg	cccatctaca	ccaacgtgac	ctatcccatt	acggtcaatc	cggcgtttgt	1800
tcccacggag	aatccgacgg	gttgttactc	gctcacattt	aatgttgtatg	aaagctggct	1860
acaggaaggc	cagacgcgaa	ttattttga	tggcgtaaac	tcggcgtttc	atctgtggtg	1920
caacgggcgc	tgggtcggtt	acggccagga	cagtcgtttt	ccgtctgaat	ttgactctgag	1980
cgcatttta	cgcgccggag	aaaaccgcct	cgcgtgtatc	gtgctgcgt	ggagtgcacgg	2040
cagtatctg	gaagatcagg	atatgtggcg	gatgagcggc	atttccgtg	acgtctcggt	2100
gctgcataaa	cogactacac	aaatcagcga	tttcatcatgtt	gcaactcgct	ttaatgatga	2160
tttcagccgc	gtctgtactgg	aggctgaagt	tcagatgtgc	ggcgagttgc	gtgactacct	2220
acgggtaaca	gtttctttat	ggcagggtga	aacgcaggtc	gccagcggca	ccgcgccttt	2280
cggcggtgaa	attatcgatg	agcgtggtgg	ttatgcccgt	cgcgtcacac	tacgtctgaa	2340
cgtcgaaaaac	cggaaactgt	ggagcggccga	aatcccgaat	cttatacggt	cggtggttgc	2400
actgcacacc	gccgacggca	cgctgattga	agcagaagcc	tgcgatgtcg	gtttccgcga	2460
ggtgcggtatt	gaaaatggtc	tgctgctgt	gaacggcaag	ccggtgctga	ttcgaggcgt	2520
taaccgtcac	gagcatcatc	ctctgcatgg	tcaggtcatg	gatgagcaga	cgtgtgtcga	2580
ggatattctg	ctgtatgaagc	agaacaactt	taacgcccgt	cgcgtttcgc	attatccgaa	2640
ccatccctg	tgttacacgc	tgtgcgaccg	ctacggcctg	tatgtggtgg	atgaagccaa	2700
tattgaaacc	cacggcatgg	tgccaatgaa	tcgtctgacc	gatgatccgc	gtggcttacc	2760
ggcgatgagc	gaacgcgtaa	cgcgaatgg	gcagcgcgtat	cgttaatcacc	cgagtggtat	2820
catctgtcg	ctggggatg	aatcaggcca	ccgcgctaatt	cacgacgcgc	tgtatcgctg	2880
gatcaaatct	gtcgatcctt	cccggccgt	gcagttatgaa	ggcggcgag	ccgacaccac	2940
ggccaccgat	attatttgc	cgatgtacgc	gcgcgtggat	gaagaccagc	ccttcccgcc	3000
tgtgcccggaa	tggtccatca	aaaaatggct	ttcgcttacct	ggagagacgc	gcccgttat	3060
ccttgcgaa	tacgcccacg	cgatggtaa	cagtcttggc	gttggcgcta	aataactggca	3120
ggcgttgcgt	cagtatcccc	gtttacaggg	cggcttcgtc	tggactggg	tggatcgtc	3180
gctgataaa	tatgtgaaa	acggcaaccc	gtggtcggct	tacggcggtg	attttggcga	3240
tacgccgaaac	gatcgccagt	tctgtatgaa	cggctctggc	tttgcgcacc	gcacccgcga	3300
tccagcgtcg	acggaagcaa	aacaccagca	gcagttttt	cagttccgtt	tatccggcga	3360
aaccatcgaa	gtgaccagcg	aatacctgtt	ccgtcatagc	gataacgagc	tcctgcactg	3420
gatggggcgc	ctggatggta	agccgctggc	aagcggtgaa	gtgcctctgg	atgtcgctcc	3480
acaaggtaaa	cagttgattt	aactgcctga	actaccgcag	ccggagagcg	ccggcaact	3540
ctggctcaca	gtacgcgtag	tgcaaccgaa	cgcgaccgca	tggtcagaag	ccgggcacat	3600
cagcgcttgg	cagcagtggc	gtctggcgga	aaacctcaatg	gtgacgctcc	ccgcccgcgtc	3660
ccacgcctac	ccgcacatctg	ccaccagcga	aatggatttt	tgcatcgagc	tggtaataaa	3720
gcgttggcaa	ttttaaccggc	agtcaggctt	tctttcacag	atgtggattt	gcpataaaaa	3780
acaactgtcg	acggcgtcgc	gcatcgttgc	caccgtgtc	gatagatctg	aacagaaact	3840
catttccgaa	gaagacctag	tgcaccatca	tcatcatcat	caccgttaat	aatagtgata	3900
taagtgtact	attagatgca	tttcgactag	atcccctgcac	caattccgg	tatttccac	3960
catatttccg	tcttttggca	atgtgaggggc	ccggaaacct	ggccctgtct	tcttgacgag	4020
cattccttagg	ggtcttcc	ctctcgccaa	aggaatgca	ggtctgttga	atgtcgtaaa	4080
ggaagcagtt	cctctggaa	cttcttggaa	acaacaaacg	tctgttagcga	cccttgcag	4140
gcagcgaac	cccccacctg	gcgacaggtg	cctctgcggc	caaaaagccac	gtgtataaga	4200
tacacctgca	aaggcggcac	aaccccagt	ccacgttgcgt	agttggatag	ttgtgaaag	4260
agtcaaatgg	ctctcctcaa	gcgttattca	caagggcgtg	aaggatgccc	agaaggtacc	4320
ccattgtatg	ggatctgatc	tggggcctcg	gtgcacatgc	tttacatgtg	tttagtcgag	4380
gttaaaaaac	gtctaggccc	cccgaaaccac	ggggacgtgg	ttttccttttgc	aaaaacacgaa	4440
tgataatacc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcg	4500
aaagttcgac	acgcgtctcc	acgtatgca	gctctcgag	ggcgaagaat	ctcggtctt	4560
cagcttcgtat	gtaggaggcgc	gtggatatgt	cctgcgggt	aataactcg	ccgatggtt	4620

ctacaaagat cgttatgttt atcggcactt tgcatcgcc ggcgtcccgaa ttccggaagt 4680  
gcttgacatt ggggaattta gcgagagcct gacctattgc atctcccgcc gtgcacaggg 4740  
tgtcacgtt caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 4800  
ggccatggat gcgatcgct cgccgatct tagccagacg agcgggttcg gcccattcg 4860  
accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 4920  
ccatgtgtat cactggcaaa ctgtatggaa cgacaccgtc aatgcgcga ttgcgcaggc 4980  
tctcgatgag ctgatgctt gggccgagga ctgccccgaa gtcgggcacc tcgtgcacgc 5040  
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 5100  
gagcgaggcg atggtcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 5160  
gtgggtggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 5220  
aggatcggcg cggctccggg cgtatatgct ccgcatttgtt cttgaccaac tctatcagag 5280  
cttgggtgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 5340  
ccgatccgga gccgggactg tcggcgta tcacatcgcc cgcagaagcg cggccgtctg 5400  
gaccgatggc tggtagaaag tactcgccga tagtgaaac cgcacgccccca gcactcgatcc 5460  
gagggcaaag gaatagagta gatggcgacc gggatctatc gataaaataa aagattttat 5520  
tttagtctcca gaaaaagggg ggaatgaaag accccacctg taggttggc aagctagctt 5580  
aagtaacgcc attttgcag gcatggaaaa atacataact gagaatagag aagttcagat 5640  
caaggtcagg aacagatgga acagctgaat atgggccaaa caggatctatc gtggtaagca 5700  
gttcctgccc cggctcaggg ccaagaacag atgaaacagc tgaatatggg ccaaacagga 5760  
tatctgttgt aagcagttcc tgccccggct caggggcaag aacagatgtt cccagatgc 5820  
ggtccagccc ttagcagttt ctagagaacc atcagatgtt tccagggtgc ccaaggacc 5880  
tgaaatgacc ctgtgcctt tttagactaa ccaatcgatc cgcttcgtc ttctgttcgc 5940  
gogcttctgc tccccggact caataaaaga gcccacaacc cctcaatcg cgcgcagtc 6000  
ctccgattga ctgagtcggc cgggtacccg tgtatccat aaacccttt gcagttgcatt 6060  
ccgacttgtt gttcgtgtt tccttggag ggtctctt gagtattgtt ctaccgtca 6120  
gccccggctt ttcattcatg cagcatgtat caaaaattaaat ttgtttttt ttcttaagta 6180  
tttacattaa atggccatag ttgcattaaat gaaatcgcca acgcgcgggg agaggcggtt 6240  
tgcgtatttg cgctttccg cttctcgct cactgactcg ctgcgcgtc ttgttcggct 6300  
gcccggagcg gtagcagtc actcaaaggc ggtataacgg ttatccacag aatcagggg 6360  
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 6420  
cgcgttgcg gcgtttttcc ataggctccg ccccccgtac gagcatcaca aaaatcgacg 6480  
ctcaagttag agtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg 6540  
aagctccctc gtgcgtctc ctgttccgac cctggccctt accggataacc tgccgcctt 6600  
tctccctcg gaaagcgtgg cgcttctca tagctcacgc ttaggttatac tcagttcggt 6660  
gtaggtcggt cgctccaagc tgggtgtgt gcacgaaccc cccgttcagc cgcacgcgt 6720  
cgccttatcc ggttaactatc gtcttgagtc caaccggta agacacgact tatcgccact 6780  
ggcagcagcc actggtaaca ggatttagcag agcgaggtat gttagcgggt ctacagagtt 6840  
cttgaagtgg tggcctaact acggctacac tagaagaaca gtattggta tctgcgtct 6900  
gctgaagcca gttacattc gaaaaagagt tggtagctt ttagccggca aacaaaccac 6960  
cgctggtagc ggtggtttt ttgtttgca gtcagcagatt acgcgcagaa aaaaaggatc 7020  
tcaagaagat ccttgatct ttctacggg gtctgacgtc cagtggaaac aaaactcagc 7080  
ttaagggatt ttggtcatga gattatcaaa aaggatctc acctagatcc ttttaaatta 7140  
aaaatgaagt ttgcggccgc aaatcaatct aaagtatata tgtagtaact tggctgaca 7200  
gttaccaatg ctaatcagt gaggcaccta tctcagcgtat ctgcgttattt cgccatcc 7260  
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccacatggcc 7320  
ccagtgctgc aatgataccg cgagacccac gtcaccggc tccagattta tcagaataa 7380  
accagccagc cggaaaggccc gagcgcagaa gttgtccgtc aactttatcc gcctccatcc 7440  
agtctattaa ttgttgcgg gaaagctagag taagtagttc gcaagttatc agtttgcgca 7500  
acgttggttgc cattgtcaca ggcatcggtt tgtagcgtc gtgttttgtt atggcttcat 7560  
tcagctccgg tttccaaacga tcaaggcggc gtcacgttcc gtcacgtc ttacatgtt 7620  
cggttagctc cttcggtct ccgatcggtt tcatgggtat ggcagactc cataattctc 7680  
tcatgggtat ggcagactc cataattctc tctgagaata gtgtatgcg agatgtttt 7740  
ctgtgactgg tgactacta accaagtcat cccgcacca tagcagaact cgaccgagg 7800  
gtcttgcgg ggcgtcaata cgggataata aactctcaag gatcttacccg tttttttttt 7860  
tcatcattgg aaaacgttct tcggggcgaa actgatette agatctttt acittccacca 7920  
ccagttcgat gtaaccctact cgtgcacccca aaaaatcgccg aaaaaaggga ataaggcg 8040  
gcgtttctgg gtgagcaaaa acagagaaggc tttttcaata ttattgaagc atttatcagg 8100  
cacggaaatg ttgaataactc atactttcc aatgtattta gaaaaataaaa caaatagggg 8160  
gttattgtct catgagcgga tacatatttg tttttttttt acitccacca 8175

E<sup>1</sup>  
cont

<210> 5  
 <211> 8175  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Nucleotide sequence for pICAST OMN

<400> 5  
 ctgcagccctg aatatggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca 60  
 gggccaaagaa cagatggAAC agctgaatat gggccaaaca ggatatctgt gttaAGCAGT 120  
 tcctgccccg gctcagggcc aagaacAGAT ggtccccaga tgcggTCCAG ccctcAGCAG 180  
 tttctagaga accatcAGAT gtttccAGGG tgccccAAgg acctgaaATG accctgtGCC 240  
 ttatTTGAAC taACCAATCA gttcgCTCTC CGCTTCTGTT CGCGCCTTC tgctccccGA 300  
 gctcaataAAAG agAGCCcACA ACCCCTCACT CGGGGCGCA gtcctCCGAT tgactgAGTC 360  
 gcccgggtac ccgtgtatcc aataaACCTC CTTGAGTGT tgactACCCG tcAGGGGGG tcttcatt 420  
 tgttccttgg gagGGTCTCC tctgAGTGT tgactACCCG tcAGGGGGG tcttcatt 480  
 gggggctcgt ccgggatcgg gagACCCCTG CCCAGGGACC accgACCCAC caccGGGAGG 540  
 caagctggcc agcaacttat ctgtgtCTGT CGATTGTCt agtgtCTATG actgatttTA 600  
 tgcgcctgcg tcggTACTAG ttagCTAAct agctCTGTat CTGGCGGACC CGTGGTGGAA 660  
 ctgacgaggT ctgaACACCC ggCCGCAACC CTGGGAGACG TCCCAGGGAC tttggggGCC 720  
 gttttgtgg cccgACCTGA ggaAGGGAGT CGATGTGGAA tccgACCCCG tcaggatATG 780  
 tggTCTGTt aggAGACGAG aacCTAAAC agtTCCGcc tccgTCTGAA ttttTGTt 840  
 cggTTTggAA ccgaAGCCGc CGTCTTGTc tgctcAGCA tCGTTCTGT ttgtCTCTGT 900  
 ctgactgtgt ttctgtattt gtctgAAAAT tagggCCAGA CTGTTACAC tCCCTTAAGT 960  
 ttgaccttag gtaACTGGAA agatgtcGAG CGGCTCGCTC acaACCAgTC ggtAGATGTC 1020  
 aagaAGAGAGC gttgggttac ctTCTGCTCT gcaGAATGGC caacCTTAA CGTCGGATGG 1080  
 cCGCgAGACG gcACCTTAA CGAGACCTC atCACCCAGG ttaAGATCAA ggtCTTTCA 1140  
 CCTGGCCCGC atggACACCC AGACCAGGT CCCTACATCG tgacCTGGGAGC AGCCTTGCT 1200  
 tttgacCCCC CTCCTGGGT caAGCCCTT GTACACCCtA AGCCTCCGC tCCCTCTCT 1260  
 ccatCCGCCC CGTCTCTCCC CCTGAACT CCTCGTTGA CCCCCTCGT atCCCTCCCT 1320  
 tATCCAGCCC tcactCCtC tCTAGGCGCC ggCCGCTCTA gCCCTTAAT AGCACTCACT 1380  
 atAGGGCGAT tcgaATCAGG CTTGGCGC CGGATCCTT aATTAAGCGC aATTGGGAGG 1440  
 tggcggTAGC CTCGAGATGG GCGTGAATT GGATTCACTG gCCGTCGTT tacaACGTG 1500  
 tgactGGGAA AACCTGGCG CTCCTTAAT TAATCGCCTT gCAGCACATC CCCCCTTCG 1560  
 cagCTGGCGT aATAGCGAAG AGGCCCCCAC CGATCGCCCT TCCCAACAGT TACCGAGCCT 1620  
 gaATGGCGAA TGGCGCTTT CCGGTTTCC GGCACCAAGAA GCGGTGCCGG AAAGCTGGCT 1680  
 ggAGTGCAT CTCCTGAGG CGATACTGT CGTCGCTCCC TCAAACCTGGC AGATGCACGG 1740  
 ttacgatGCG CCCATCTACA CCAACGTGAC CTATCCATT ACGGTCAATC CGCCGTTGT 1800  
 tcccacGGAG aATCCGACGG GTTGTACTC GCTCACATT AATGTTGATG AAAGCTGGCT 1860  
 acAGGAAGGC CAGACGCGAA TTATTTTGA TGGCGTTAAC TCGCGTTTC ATCTGTGGGT 1920  
 caacGGGCGC TGGGTCGGT ACGGCCAGGA CAGTCGTTG CGTCTGAAT TTGACCTGAG 1980  
 CGCATTtTA CGCGCCGGAG AAAACCGCT CGCGGTGATG GTGCTGCGCT GGAGTGCAGG 2040  
 cagttatCTG gaAGATCAGG GCTGATACTC GTCACATTT AATGTTGATG AAAGCTGGCT 2100  
 gctgcataAAAG CCGACTACAC AAATCAGCGA TTTCCATGTT GCCACTCGCT TTAATGATGA 2160  
 tttcAGCCGc GCTGTACTGG AGGCTGAAGT TCAAGATGTGc GGCAGGTTGc GTGACTACCT 2220  
 acGGGTAAcA GTTTCTTTAT GGCAGGGTGA AACCGAGGTC GCCAGCGGA CGCGCCCTT 2280  
 CGGCGGTGAA ATTATCGATG AGCGTGGTGG TTAGGCCAT CGCGTCACAC TACGTCTGAA 2340  
 CGTCGAAACAC CGAAACTGT GGAGCGCCGA AATCCCGAAT CTCTATCGT CGGTGGTTGA 2400  
 ACTGCACACC GCGACCGCA CGCTGATTGA AGCAGAACCC TGCAGATGTG CGTCTCCGCA 2460  
 ggtgcggatt gaaaATGGTC TGCTGCTGT GAACGGCAAG CGTCTGAAT TTGAGGGCGT 2520  
 taaccgtcac gagcatcatc CTCTGCATGG TCAAGGTATG GATGAGCAGA CGATGGTGC 2580  
 ggatATCCTG CTGATGAAGC AGAACAACTT TAACGCCGTG CGCTGTTGC ATTATCCGAA 2640  
 CCATCCGCTG TGGTACACGC TGTGCAGCGC CTACGGCCTG TATGTGGTGG ATGAAGCCAA 2700  
 tattGAAACC CACGCGATGG TGCCAATGAA TCGTCTGACC GATGATCCGc GCTGGCTACC 2760  
 ggcgatGAGC gAACCGTAA CGCGAATGGT CGAGCGCAT CGTAATCACC CGAGTGTGAT 2820  
 catCTGGTCG CTGGGAATG aATCAGGGCA CGCGCTAAT CACGACGCGC TGTATCGCTG 2880  
 gatcaaatCT GTCGATCCTT CCCGCCGGT CGAGTATGAA GGCAGCGGAG CGACACCCAC 2940  
 ggCCACCGAT ATTATTCGCC CGATGTACGC CGCGCTGGAT GAAGACCGAGC CTTCCCGGC 3000  
 tGTGCCGAAA TGGTCCATCA AAAAATGGCT TTGCTCACCT GGAGAGACGC GCCCCTGAT 3060  
 CCTTTCGAA TACGCCACG CGATGGTAA CGTCTTGGC GGTTCGCTA AATACTGGCA 3120  
 ggcgtttcgt cagtatcccc GTTACAGGG CGGCTCGTC TGGACTGGG TGGATCAGTC 3180

6  
 Cont

gctgattaaa tatgatgaaa acggcaaccc gtggtcggct tacggcggtg attttggcga 3240  
 tacgccgaac gatgccagt tctgtatgaa cggctctggc tttgccgacc gcacgcccga 3300  
 tccagcgctg acggaagcaa aacaccagca gcagttttc cagttccgtt tatccggca 3360  
 aaccatcgaa gtgaccagcg aatacctgtt ccgtcatagc gataacgagc tcctgcactg 3420  
 gatggtggcg ctggatggta agccgctggc aagcggtgaa gtgcctctgg atgtcgctcc 3480  
 acaaggtaaa cagtgattt aactgcctga actaccgcag ccggagagcg ccgggcaact 3540  
 ctggctcaca gtacgcgtag tgcaaccgaa cgcgaccgca tggtcagaag ccgggacat 3600  
 cagcgcctgg cagcagtggc gtctgggga aaacctcagt gtgacgcctcc cgcgcgcgtc 3660  
 ccacgcctatcc cgcatcgaa ccaccagcg aatggattt tgcacgacg tggtaataa 3720  
 gcgttggcaa tttaccgcgc agtcaggctt tcttcacag atgtggattt gcgataaaaa 3780  
 acaactctg acgcgcgtgc gcgatcaat caccctgtc gatagatctg aacagaaaact 3840  
 catttccgaa gaagacctag tcgaccatca tcatcatcat caccggtaat aataggtaga 3900  
 taagtgactg attagatgca ttgcactatc atccctcgac caattccgt tattttccac 3960  
 catattgcgc tctttggca atgtgaggcc ccggaaaccc ggcctgtct tcttgacgag 4020  
 cattccttagg ggtctttccc ctctcgccaa aggaatgcaat ggtctgttga atgtcgtaa 4080  
 ggaagcgagg cctctggaa gcaagcgaaaccc cccgcgtc caaaagccac gtgtataaga 4140  
 gcaacgcgtc ccccccacccg tacacctgca aaggcgac aaccccgatg ccacgttgcg agttggatag ttgtggaaag 4260  
 agtcaaatgg ctctctcaa ccattgtatg ggatctgatc gcttggggctcg tggcataatgttggcc 4320  
 gttaaaaaac gtctaggccc cccgaaccac ggggacgtgg tttcccttg aaaaacacaga 4440  
 tgataatacc atgaaaaaaggc ctgaaactc acgcgacgtct gtcgagaagt ttctgatcg 4500  
 aaagttcgc acgcgtctccg cagttcgat gttaggaggc ctacaaagat cgttatgtt gcttggggctcg aaggatgccc aagaggtaacc 4560  
 ggtttcgatc gttatgttgc tccaaacaatg gatggggctcg atcggcaactt tgccatcgatc 4620  
 atcggcaactt tgcatcgcc ggcctcccgaa ttccggaaatgttgcgatggcc 4680  
 gcttggacattt ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 4740  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 4800  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 4860  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 4920  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 4980  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5040  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5100  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5160  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5220  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5280  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5340  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5400  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5460  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5520  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5580  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5640  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5700  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5760  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5820  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5880  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 5940  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6000  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6060  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6120  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6180  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6240  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6300  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6360  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6420  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6480  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6540  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6600  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6660  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6720  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6780  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6840  
 tgtaatggcggcc ggggaaatggc tgtaatggcggcc gcgagggctt gacctattgc atctccgcgtc gtgcacaggg 6900

1  
 Cont

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 6960  
 cgctggtagc ggtgttttt ttgttgcaa gcagcagatt acgcgcagaa aaaaaggatc 7020  
 tcaagaagat ccttgatct tttctacggg gtctgacgct cagtgaacg aaaactcacg 7080  
 ttaagggatt ttggcatga gattatcaa aaggatcttcc acctagatcc ttttaaatta 7140  
 aaaatgaagt ttgcggccgc aaatcaatct aaagtatata tgagtaaact tggctgaca 7200  
 gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 7260  
 tagttgcctg actccccgtc gtgttagataa ctacgatacg ggagggctta ccatctggcc 7320  
 ccagtgtgc aatgataccg cgagaccac gtcaccggc tccagattt tcagcaataa 7380  
 accagccagc cggaagggcc gagcgcagaa gtggcttgc aactttatcc gcctccatcc 7440  
 agtctattaa ttgttgcgg gaagctagag taatgttgc gccaggtaat agtttgcgca 7500  
 acgttggc cattgtaca ggcatcggtt tgcacgctc gtcgttggg atggcttcat 7560  
 tcagctccgg ttcccaacga tcaaggcgg tagcatgatc ccccatgttg tgcaaaaaaag 7620  
 cggtagctc cttcggtctt ccgatcggtt tcagaagtaa gttggccgca gtgttatcac 7680  
 tcatggttat ggcagactg cataatttctc ttactgtcat gccatccgta agatgcttt 7740  
 ctgtgactgg tgagactca accaagtcat tctgagaata gtgtatgcgg cgaccgagg 7800  
 gctctgccc ggcgtcaata cgggataata cccgcacca tagcagaact taaaaagtgc 7860  
 tcatcattgg aaaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 7920  
 ccagttcgat gtaacccact cgtgcacca actgatctt agcatcttt actttcacca 7980  
 gcgttctgg gtgagcaaaa acaggaaggc aaaaatgccgc aaaaaaggga ataaggcga 8040  
 cacggaaatg ttgaataactc atactcttcc ttttcaata ttattgaagc atttatcagg 8100  
 gttattgtct catgagcgg tacatattt aatgtattta gaaaaataaa caaatagggg 8160  
 ttccgcgcac atttc 8175

C1  
 <210> 6  
 <211> 5  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> GS Linker

<400> 6  
 Gly Gly Gly Gly Ser  
 1 5